

UNDERSTANDING INTERNET PERFORMANCE FROM THE USER PERSPECTIVE

Kacker, Liu, Yen, Zhang, Wilkinson,
Marbukh, Kelley, Mills, Montgomery

DARPA Network Modeling and Simulation Workshop
09/27-29/00, Albuquerque, NM

PROJECT GOALS & NMS CHALLENGES

- Develop a web-based repository of “useful real data” with interactive graphical and statistical analyses that can be extended by the DARPA NMS Research Community.
 - Addresses NMS challenge to validate models by providing real data and useful analysis routines to characterize real data.
 - Addresses NMS challenge to produce innovative models of traffic sources by focusing on analyses to produce statistic functions that represent measured traffic.
- Investigate techniques to identify and respond to anomalies in network conditions, even in the face of uncertainty regarding the available information
 - Addresses NMS challenge to provide innovative models of network control by applying game theoretic approaches to controlling network resource allocation under uncertainty.
 - Addresses NMS challenge to provide innovative models of network traffic through exploratory data analysis to produce patterns that can be used for detection of network anomalies.
- Improve existing techniques for modeling and understanding Internet behavior and performance
 - Addresses NMS challenge to provide innovative models of networks by adding features to existing DARPA-funded simulation models and by applying cellular automata models to investigate network dynamics arising from collective behavior at multiple time scales.

PROGRESS & FUTURE PLANS IN THREE AREAS

1. EXPLORATORY DATA ANALYSIS USING DATA COLLECTED DURING TWO YEARS OF XIWT ACTIVE MEASUREMENTS
2. FRAMEWORK FOR MODELING NETWORK RESOURCE ALLOCATION UNDER UNCERTAIN INFORMATION
3. SIMULATION MODELS FOR DIFFERENTIATED SERVICES QUEUE MANAGEMENT ALGORITHMS

EXPLORATORY DATA ANALYSIS

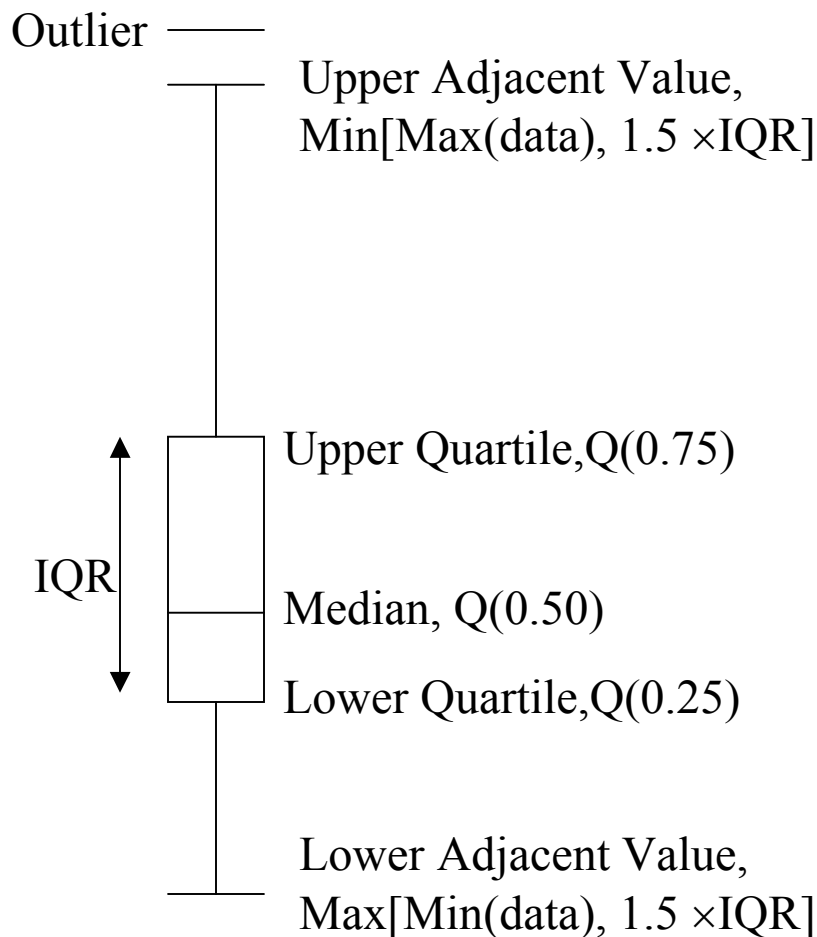
- Accomplishments
- Proof-of-Concept Interactive Data Analysis Scripts
 - Box Plots for Round-Trip Times
 - Auto-Correlation Functions for Round-Trip Times
- Future Plans

ACCOMPLISHMENTS

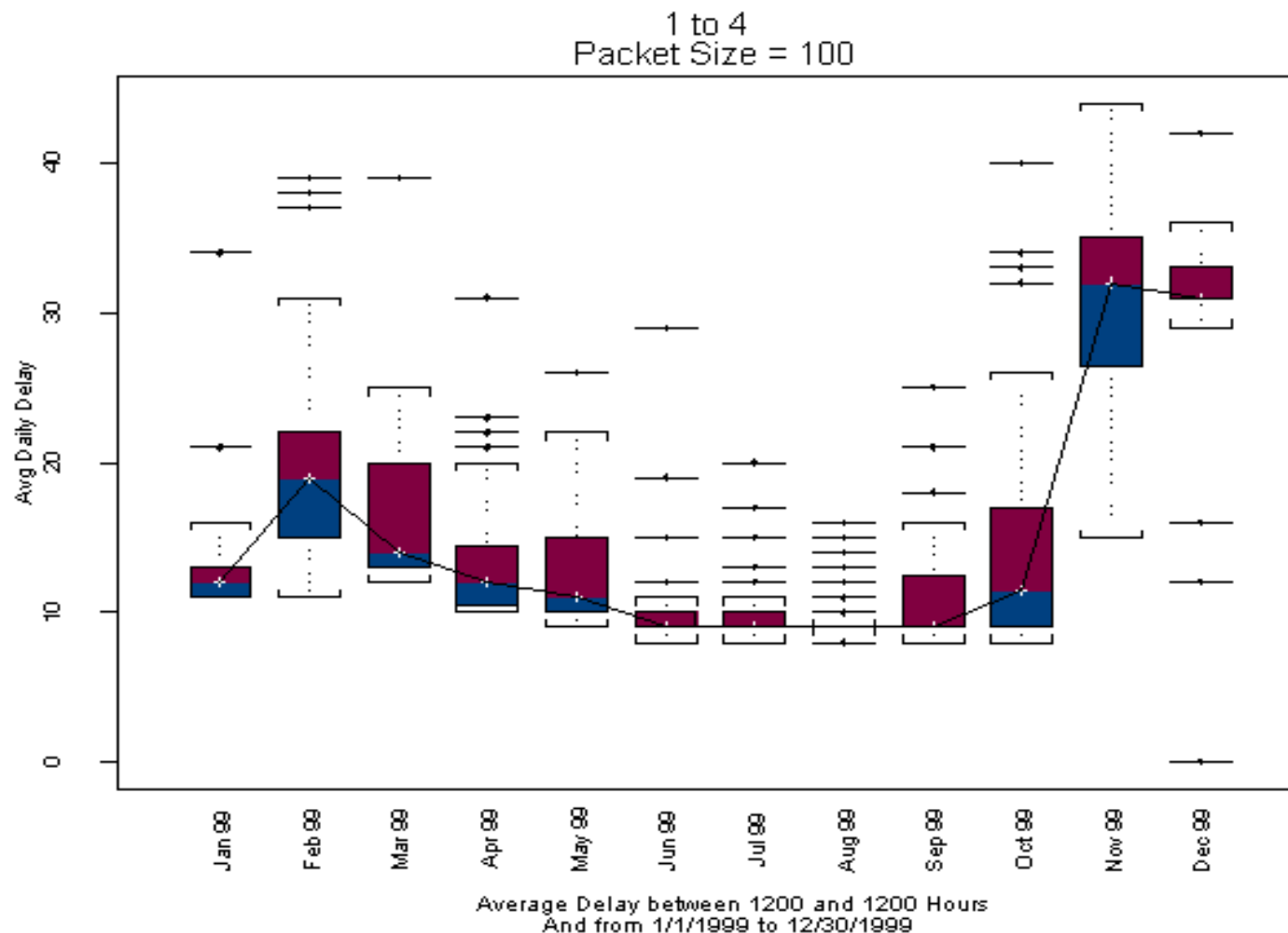
- **Proof-of-Concept**
 - Data: 05/01/1998 - 04/30/2000
 - Analysis Scripts: subset, Box plot, and Auto correlation function
 - Prototype: <http://statserver.statsci.com/statserver/demos>
 - Underway: Converting Proof-of-Concept into Operational Web-based StatServer Located at NIST
- **Exploratory Data Analysis**
 - Using Round-Trip Times from NIST to 3 destinations and covering the period from 05/01/98 - 04/09/99
- **Findings**
 - For a given time interval, stationary models apply, but different models for different time intervals
 - Can be modeled as long memory process with cyclic behavior
 - Heavy-tailed distributions

BOX PLOT

A five-point summary
of one dimensional data



A BOX PLOT FROM XIWT RTT DATA



AUTO-CORRELATION FUNCTION (ACF)

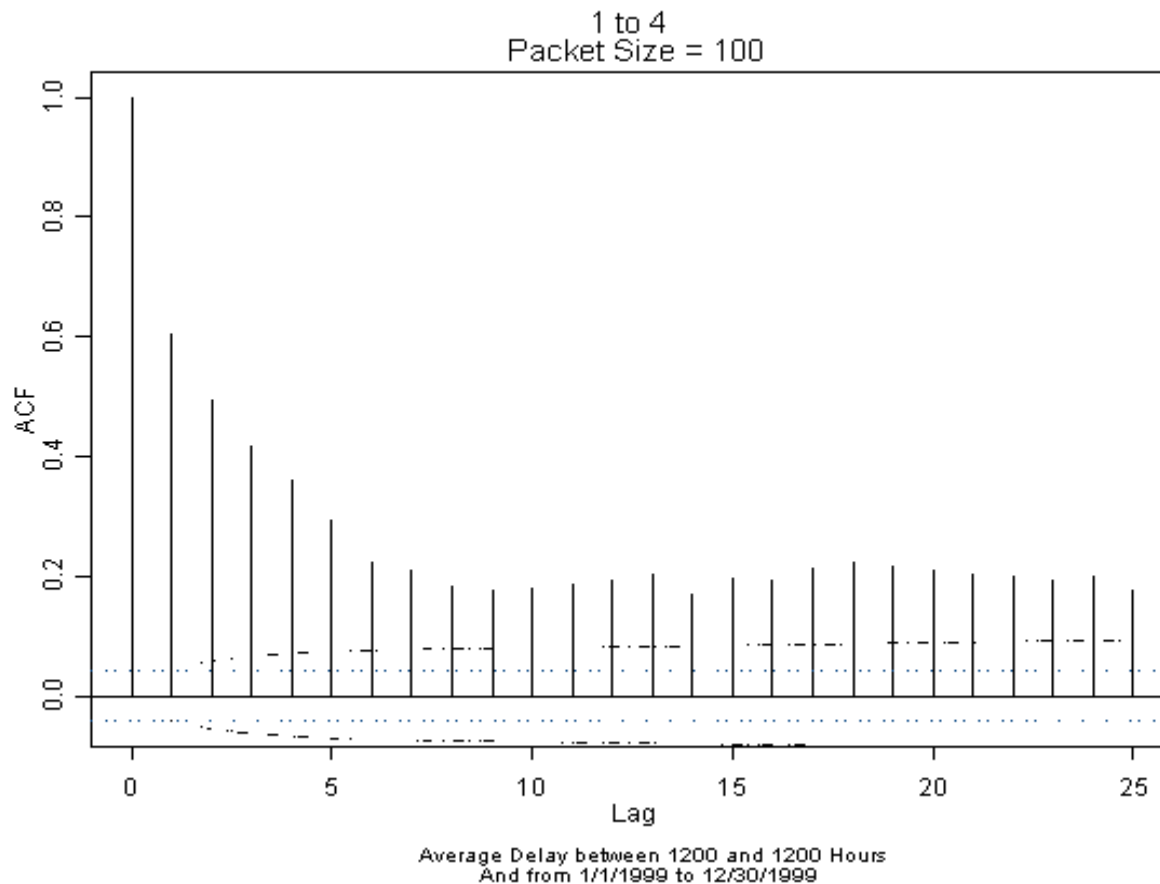
- Data Set: X_1, X_2, \dots
- ACF (Lag k): correlation between the data

– $X_1, X_2, \dots, X_k, X_{k+1}, X_{k+2}, \dots, X_N, \dots, X_{k+N}$

– $X_1, X_2, \dots, X_k, X_{k+1}, X_{k+2}, \dots, X_N, \dots, X_{k+N}$

- Represents structure in time series $X_1, X_2, \dots, X_N, \dots, X_{k+N}$
 - “near zero” for white noise process
 - “tapers off” for auto-regressive process
 - “cuts off” for moving average process
 - “remains high” for random walk

ACF PLOT FROM XIWT RTT DATA



FUTURE PLANS

- Bring web-based XIWT data and analysis service “on the air”
 - Target date: mid-January 2001
 - Still must deal with: missing data, data integrity, and maintenance
 - Automate copy of CNRI data to NIST once a day
 - Subsets of data
 - Provide baseline capabilities: XIWT Internet Service Performance, Data Analysis and Visualization, and other analytics
 - Seek feedback on what analytics are of interest and on other improvements
- Once web-based data repository with integrated statistical software is operation, we will consider expanding to other “useful data sets” and supporting analysis scripts from others

Network Modeling and Control (NM&C) under Uncertainty

- Problem Statement
- General Framework for a Solution
- Simple but Non-Trivial Example
- Future Plans

Problem Statement

- Sources of Uncertainty
 - Statistical nature of measurements, i.e., confidence intervals rather than point estimates for the parameters
 - Possible non-stationary environment, i.e., anomalies
 - Possible adversarial environment, i.e., denial of service attack
- Currently NM&C deals only with extreme cases
 - **Probabilistic approach:** assumes simple typically stationary probabilistic models, with known parameters, for the environment. Average, steady-state behavior. Very sensitive to the assumptions.
 - **Competitive approach:** makes no assumptions on the environment. Guards against the worst case scenario. Results in very expensive solutions in terms of the resource utilization.
- **Challenge:** develop middle-ground approach to NM&C able to incorporate partial information on the environment

General Framework for NM&C under Uncertainty

Step 1. Probabilistic approach

$$L^{opt}(\theta) = \min_{u \in U} L(u, \theta) \Rightarrow u^{opt}(\theta)$$

where: L -losses, u -control action, θ -environment

Step 2. *Minimax* approach

$$F_*^* = \min_{u \in U} \max_{\theta \in \Theta} \{L(u, \theta) - L^{\min}(\theta)\} \Rightarrow u^{opt}(\Theta)$$

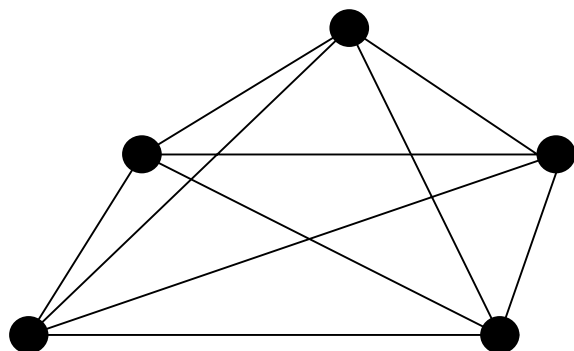
Step 3. *Minimax&Bayes'* approach

$$u \in \{u : \Phi(u) - F_*^* \leq \Delta, u \in U\}$$

where $\Phi(u, \theta) = \max_{\theta \in \Theta} F(u, \theta)$

and risk tolerance of the network is characterized by $\Delta > 0$

Example: Admission control in a symmetric fully connected circuit switched network



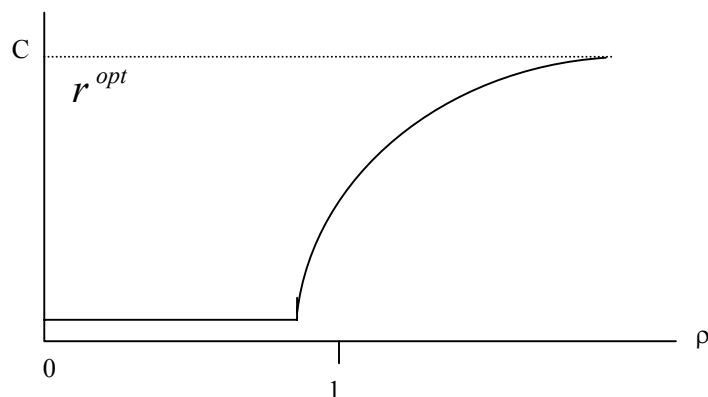
link capacity: C

link load: $\lambda = C\rho$

trunk reservation parameter: r

loss probability: $L(r, \rho)$

$$u \Rightarrow r, \theta \Rightarrow \rho$$



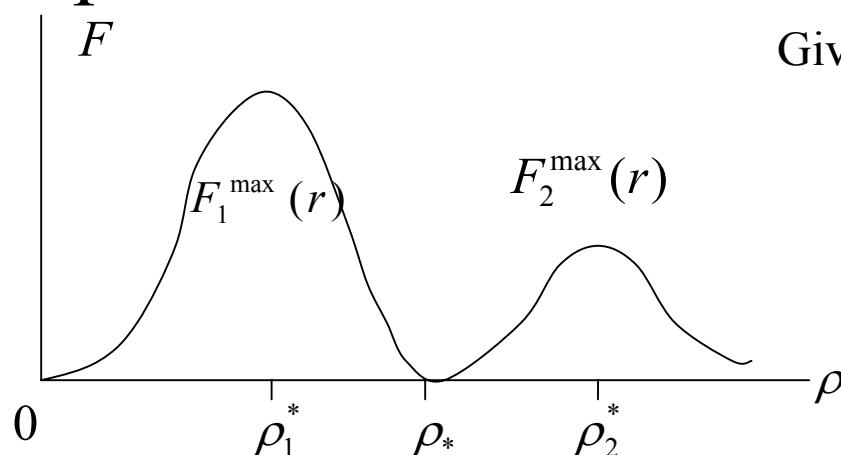
Given ρ , optimal trunk reservation:

$$L^{\min}(\rho) = \min_{r=0,1,\dots,C} L(r, \rho) \Rightarrow r^{opt}(\rho)$$

Links occupancies are independent if

$$N \rightarrow \infty$$

Example continued: solution of the *minimax* problem for the trunk reservation parameter



Given confidence interval for the external load

$$[\rho - \delta, \rho + \delta]$$

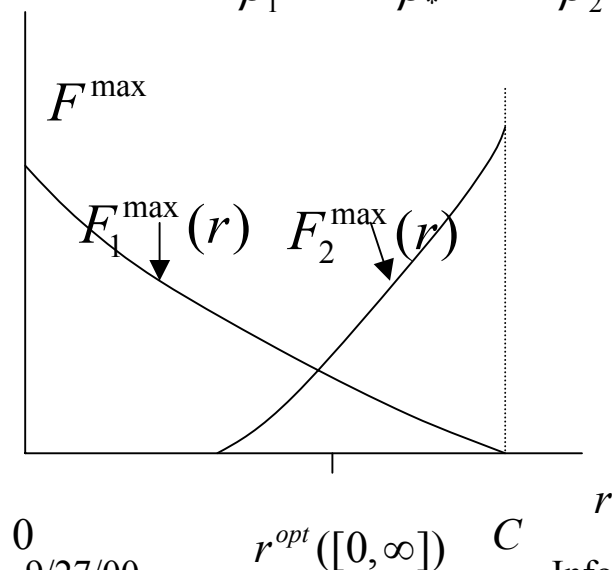
find the trunk reservation parameter

$$r^{opt}([\rho - \delta, \rho + \delta])$$

by solving optimization problem

$$\min_{r=0,1,\dots,C} \max_{\rho - \delta \leq \tilde{\rho} \leq \rho + \delta} F(r, \tilde{\rho})$$

$$\text{for } F(r, \tilde{\rho}) = L(r, \tilde{\rho}) - \min_{r=0,1,\dots,C} L(r, \tilde{\rho})$$

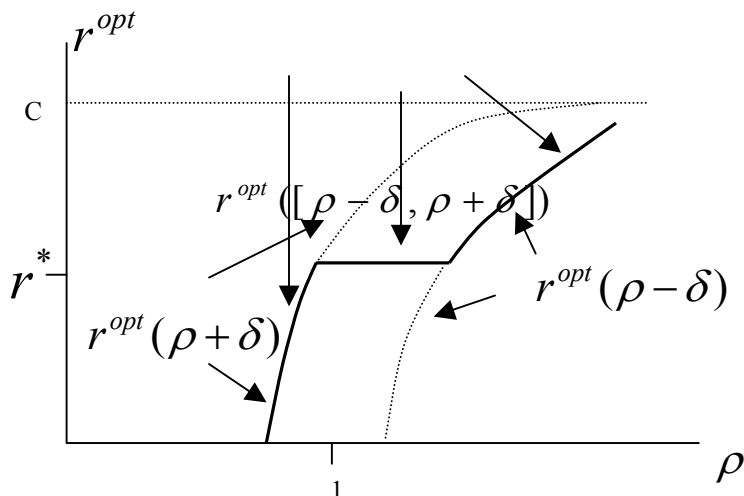


Function $F(r, \tilde{\rho}) \geq 0$ has three local minima over

$$\tilde{\rho} : F(r, \tilde{\rho}) = 0 \quad \text{at} \quad \tilde{\rho} = 0, \tilde{\rho} = \infty$$

$$\text{and} \quad \tilde{\rho} = \rho_* : r^{opt}(\rho_*) = r$$

Results for the Example and Future Research



- * If no information on ρ is available, the optimal trunk reservation parameter is

$$r^{opt}([0, \infty]) = r^*$$

Example: $C = 10 \Rightarrow r^* = 2$

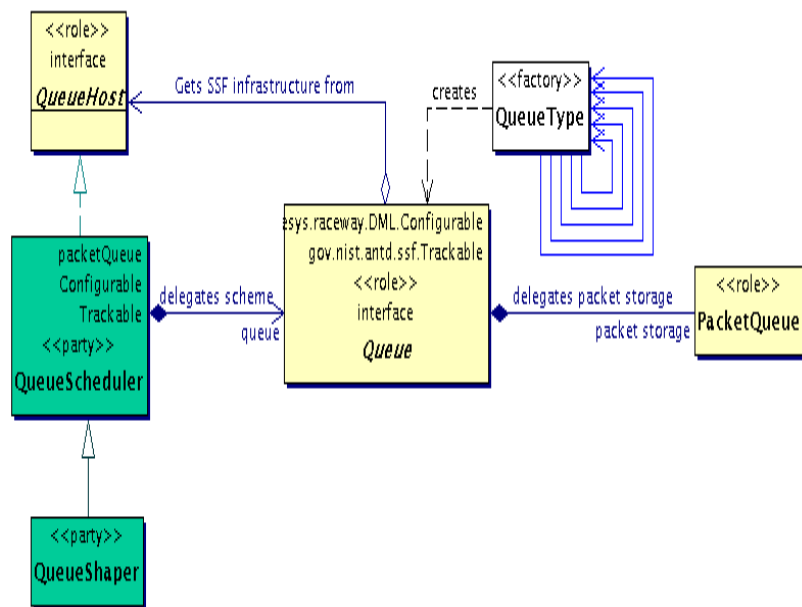
- * Solution is based on boundaries (not average) of confidence interval $[\rho - \delta, \rho + \delta]$

Future research:

- * incorporate incomplete information into network model
- * develop modeling technique for non-stationary environment
- * implement this approach for more realistic network models (Internet)
- * assess sensitivity of the network performance with respect to measurement errors

OTHER PROGRESS

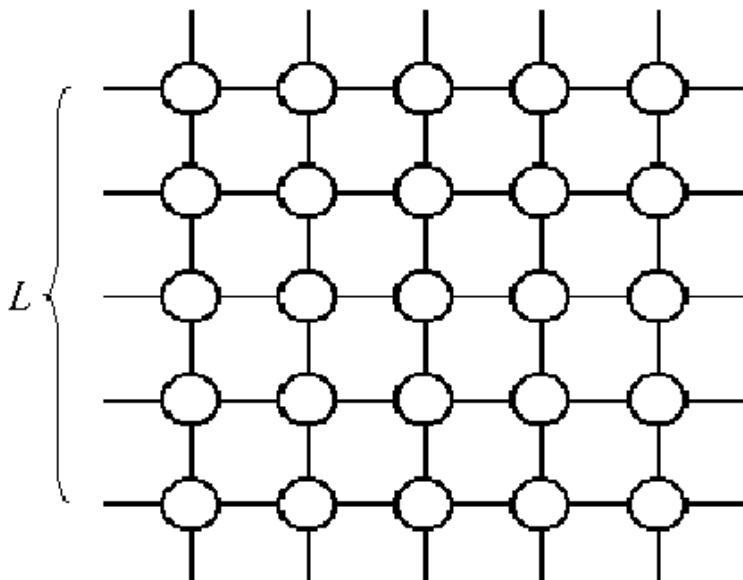
DEVELOPED MODELS FOR DIFFERENTIATED SERVICES QUEUE MANAGEMENT ALGORITHMS



CONTRIBUTED MODELS TO
FUTURE RELEASE OF SSFnet

- RedQueue - implements the RED scheme
- RioQueue - implements the RIO scheme
- MeterQueue - tags packets according to the marking scheme
- TSW2CMarker - implements the TSW algorithm to mark packets "In" or "Out"
- PremiumQueue - implements the core router part of the Premium Service

OTHER PLANS



- Investigate the application of 2-D Cellular Automata to Model Dynamics of Large-Scale Networks on Multiple Time Scales
- Explore for phenomena caused by collective behavior of autonomous interacting network nodes
- Visualize Time-Varying Dynamic Behavior of Large-Scale Networks